

Cosmic CO and [CII] backgrounds and the fueling of star formation over 12 Gyr

Corresponding Author: Dr Yi-Kuan Chiang

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 1:

Reviewer comments:

Reviewer #1

(Comments for the Author)

I enjoyed reading the paper. I want to congratulate the authors on this work. Following is my assessment with some questions and comments:

1. Summary:

The paper makes the claims of reporting the first detections of the cosmic CO background aggregate intensity (monopole) at 6.6σ and the cosmic [CII] background at 3.0σ , spanning $0 < z < 4$. The analysis is based on an application of tomographic clustering of diffuse intensity maps (Planck, Herschel, IRAS) cross-correlated with spectroscopic reference galaxies and quasars from the SDSS survey.

2. Originality and Significance:

The work represents an important advance by providing the first empirical constraint on the mean intensity (ρ_{CO}) of the CO background. This approach bypasses the complex, model-dependent integration of the CO luminosity function needed by previous power spectrum measurements (e.g., Keating et al. 2020), offering a crucial normalization factor for the cosmic molecular gas reservoir. The finding that the derived Ω_{H_2} is about twice that of galaxy surveys will directly influence models of galaxy evolution. If the robustness of these results is established, it provides an anchor point for the molecular gas content of the universe, offering a benchmark against which all future theoretical models (e.g., IllustrisTNG, EAGLE) and LIM experiments can be validated. The resulting empirical constraints on the Kennicutt-Schmidt law across cosmic time are important too.

3. Critical Assessment and Comments:

I think the work is largely convincing but is dependent on a chain of non-trivial model assumptions that obscure the physical interpretation of the final astrophysical quantities (Ω_{H_2} , Σ_{SFR}). To fully strengthen the conclusions and make the results more robust, further clarifications are required in the following areas:

MAJOR COMMENTS:

i) Model Dependence and α_{CO} (claims of a "Direct Measurement"):

(Lines 85, 92): The manuscript must clearly distinguish between a 'measurement' and an 'inference' based on model assumptions. The final determination of the molecular gas density Ω_{H_2} is an inferred value, not a "direct measurement". It relies on applying a fixed, non-universal α_{CO} conversion factor. This places the inherent process of this paper in the same spirit as other LIM results (e.g., Breyse et al. 2021) which estimate ρ_{H_2} based on CO intensity, modulated by conversion assumptions. Perhaps the text must be rephrased to emphasize that this is the first unbiased constraint on ρ_{CO} monopole, and the resulting Ω_{H_2} is an inference using the assumed α_{CO} .

(Lines 95-97): The claim that Ω_{H_2} is twice that of galaxy surveys significantly challenges the models of galaxy evolution. It relies on this same assumption about α_{CO} . α_{CO} is known to vary significantly with metallicity and gas dynamics (e.g. Bolatto et al. 2013), and the aggregate cosmic population averages over these evolving properties. As a result, this comparison is potentially biased if the galaxy surveys targeted high-metallicity galaxies (low α_{CO}) while the cosmic average requires a different, possibly higher, α_{CO} due to the inclusion of metal-poor populations (Tacconi

et al. 2020). It would be good to see an explicit discussion on how the factor-of-two discrepancy would change if a realistic, redshift-dependent $\alpha_{\text{CO}}(z)$ (e.g., scaling with average cosmic metallicity) were adopted. Optionally, the authors can quantify the systematic uncertainty on Ω_{H2} if α_{CO} were allowed to vary within a conservative, physically plausible range (e.g., 0.8 to 5.0) across $0 < z < 4$. The text must be rephrased to state that the value is inferred via an assumed α_{CO} value.

ii) Low- z CIB-Line Degeneracy:

(Equation 10-12): The [CII] is unconstrained at $z < 0.4$ due to a gap in spectral coverage. This means the [CII] amplitude at $z=0$ ($R_{\text{[CII]},0}^{100}$) is poorly constrained by the data. The fit from authors uses an evolution parameter, $C_{\text{[CII]}^{100}}$, which dictates how the [CII] line-to-continuum ratio changes with z (Equation 10-12). This parameter is weakly constrained by the data, reflected by its large posterior error (Extended Data Table 2). The model determines the [CII] cosmic history by simultaneously solving for this $z=0$ anchor ($R_{\text{[CII]},0}^{100}$) and its evolution ($C_{\text{[CII]}^{100}}$) over the range where it is observed ($z > 0.4$). Since the data cannot pin down the $z=0$ anchor point, the fit must extrapolate the CIB continuum (which defines the [CII] baseline) and the evolution of the line down to $z=0$. This is difficult because the weakly constrained ($C_{\text{[CII]}^{100}}$) parameter directly governs the slope of the [CII] history (Fig. 1b). If the model inaccurately estimates the [CII] normalization at $z=0$ (due to weak constraints), this error is carried forward and propagates across all redshifts through the evolution parameter, $C_{\text{[CII]}^{100}}$. A small error in the assumed baseline or slope can significantly change the inferred ρ_{CII} at the peak star formation epoch ($z \sim 2$). The sensitivity is further complicated because at low frequencies/low z , CO lines are most prominent (Fig. 4). Although the CO and [CII] are distinct lines, the spectral fitting process is a complex multi-parameter process; if one part of the low- z spectrum is uncertain due to [CII] extrapolation, it could subtly influence the broader continuum or other co-existing line components.

Have any robustness tests been done to check how sensitive are the high-redshift ρ_{CII} results are to the extrapolation of the CIB continuum model at low frequencies? Have the authors performed a fit using some alternative, motivated prior on the [CII] evolution, such as a fixed (non-evolving) ratio (e.g. $C_{\text{[CII]}^{100}} = 0$, enforcing a flat relationship with the continuum), or a redshift-anchored power law form for $R_{\text{[CII]}}(z)$. If the peak ρ_{CII} result at $z \sim 2$ (where coverage is complete) is truly driven by the data, it should remain stable regardless of the form used for the low- z extrapolation. Forcing $C_{\text{[CII]}^{100}} = 0$ can test whether the weak prior $N(0, 2)$ can artificially steer the high- z results. These tests are just suggestions, but I think the authors should comment on the robustness of this method since it is an important part of claiming the detection and they would add significant credibility to the detection.

iii) CO SLED Expectation:

(Line 242): The authors claim that the SLED should be "representative of the latter [main-sequence galaxies]". This is an assumption that risks biasing the inferred Σ_{SFR} . The aggregate signal includes all types of galaxies including the luminous starbursts which have highly excited SLEDs. This could be an important caveat because the CO SLED model is calibrated based on a small sample of CO-detected star-forming galaxies at $z=1-3$. This risks introducing a bias based on the properties of those typically bright, resolved sources. Is this a worry? The authors should justify this statement with some prior reference or through some expectations for contribution of main-sequence vs. starburst galaxies to the overall ρ_{TIR} and ρ_{CO} . Will the results on Σ_{SFR} and the steepness of the Kennicutt-Schmidt (K-S) law ($N \sim 1.5$) change significantly based on the scatter in the observed galaxy SLEDs and/or specific choice of the CO SLED function? I don't think a detailed discussion is required here, just some small justification should suffice.

MINOR COMMENTS:

iv) K-S Law Surface vs. Volume Densities:

(Lines 131-140): The analysis linking ρ_{H2} and ρ_{SFR} (volume) to derive t_{dep} , and then using t_{dep} to calculate Σ_{H2} and Σ_{SFR} (surface), relies on the assumption that the cosmic-mean depletion time is equal to the average galactic depletion time. This leap from a volume average to a surface average is a central, and slightly non-trivial assumption. It would great if the authors could state this core assumption in the text.

Along the same lines, I have two more curiosities. These following two are optional and a discussion is not needed for the paper:

a) results show that Σ_{SFR} rises with redshift (Fig. 3b). Given that t_{dep} decreases with redshift (Fig. 3c), what is the physical driver here? Does the average star-forming galaxy at high z simply have a smaller star-forming disk size for a fixed stellar mass, or is this primarily a consequence of higher gas fractions/turbulent pressures?

b) implications for Semi-Empirical Models: Fig. 2 shows the Ω_{H2} results are in better agreement with the EAGLE and semi-empirical model predictions than with IllustrisTNG. What specific physical or empirical parameter in the EAGLE or semi-empirical models (e.g., the halo-mass-to-CO luminosity relation, or the CO-to-H2 conversion) leads to their higher Ω_{H2} that aligns better with this novel LIM measurement?

v) Thermal Sunyaev-Zel'dovich effect (tSZ) and radio contamination:

The Planck maps contain tSZ (100, 143, and 353 GHz; although smaller than CIB at 353 GHz) and radio emission from extragalactic galaxies (mostly at 100 and 143 GHz). Since these emissions also trace the large-scale structure, they should be correlated with the galaxies and quasars from the SDSS galaxy surveys used here. While the companion paper mentions fitting for a separate tSZ term, no such thing has been mentioned here. The authors should clarify how this has been dealt with. Also, the authors should specify if they expect the radio contamination to be negligible if it has been neglected.

vi) Neglecting Higher-J CO lines and Interlopers:

(Lines 55-56): The model explicitly stops fitting the CO ladder at $J=9-8$. High-J transitions ($J \geq 10$) are powerful tracers of the dense, highly excited gas that characterizes starbursts (e.g. Greve et al. 2014) and should contribute to the observed bands used here. Ignoring these ignores a fraction of the total luminosity and potentially changes the inferred Σ_{SFR} based on SLED modeling.

At the same time, the paper neglects other bright atomic lines ([OIII], [CII]) and dense gas tracers (HCN, HCO+) that correlate with star formation and fall within the observational windows. If these are expected to be negligible, it would make the paper more thorough to mention this justification for the exclusion of $J \geq 10$ lines as well as other interloper lines.

vii) SLED Ceiling:

(Line 251): Provide a literature reference or argument for the fully thermalized ceiling assumption ($f_{\text{SLED}} = J^2$ at $\log \Sigma_{\text{SFR}} = 4$).

viii) CONCERTO Reference:

(Line 365): CONCERTO is not a model. Correct the phrasing to confirm that Ref. 64 (B  thermin et al. 2022) is cited as a simulation/model (a forecast for line emissions) that produces results consistent with the detection, rather than referring to the observational experiment.

ix) CO First Detection:

In the abstract, the authors mention that this is the first detection of the cosmic CO background. While the distinction between measuring the monopole (mean intensity) and the power spectrum (shot noise) is key, the term first detection can be ambiguous given previous LIM work like COPSS and mmIME. It would be good to clearly state that this first detection refers specifically to the mean intensity (ρ_{CO}), which is the integral (first moment) of the luminosity function, as opposed to the power spectrum (second moment) detected previously.

x) ρ_{SFR} Consistency and Possible Circular Logic:

(Lines 121-123): The dependence of the depletion time (t_{dep}) on $\rho_{\text{SFR}}(z)$ (Line 122), which references the CIB-only companion study, could be problematic because that original model neglected line contributions. While the present joint fit seems to correct the biased CIB parameters (Σ_{beta} changes from 0.71 to 0.24), the language suggests using the old result from the companion paper. Can the authors please explicitly confirm in the Methods section that the $\rho_{\text{SFR}}(z)$ used for t_{dep} is derived from the corrected CIB continuum parameters obtained from the joint CIB+lines fit performed in this current paper or the old paper?

Congratulations on the work once again,
Abhishek S. Maniyar

(Remarks on code availability)
NA

Reviewer #2

(Comments for the Author)
Key Results:

This manuscript presents the first tomographic detection of the cosmic CO and [CII] backgrounds over $0 < z < 4$ using cross-correlations between broadband far-infrared/millimeter intensity maps (Planck, Herschel, IRAS) and spectroscopic tracers from SDSS. The authors recover redshift-resolved, bias-weighted emissivities and forward-model the continuum and line components using a 21-parameter framework combining a graybody CIB model with a CO SLED and a [CII] component. They infer a molecular gas density approximately two times higher than galaxy survey estimates at $1 < z < 2.5$, a depletion time of 1 Gyr, and a persistent large-scale Kennicutt–Schmidt relation. These results provide an important empirical baseline for future LIM experiments.

Validity:

I do not find problems that would preclude publication, but I do identify important points that require clarification, correction or expansion before the results can be considered robust:

1) There appears to be a significant inconsistency between Equation (6), Extended Data Table 1, and Figure 3a. This may be a definitional or scaling issue, but it requires clarification or correction.

For example:

For $J=2-1$, at $\log \Sigma_{\text{SFR}} = 0$ (i.e., 1 Msun/yr/kpc^2):

$\log f_{\text{SLED}} = 0.3 \times (0 - (-2))^{1.95} + 0.31 = 1.47$

So: $I(2-1)/I(1-0) = 10^{1.47} \sim 30$

But Fig 3a shows at $J_{\text{up}}=2$ the cosmic ratios are $\sim 0.5-3$.

The text also states fSLED should reach J^2 at $\log \Sigma(\text{SFR}) = 4$.

For $J=2-1$ at this $\Sigma(\text{SFR})$ value:

$\log f_{\text{SLED}} = 0.3 \times (4 - (-2))^{1.95} + 0.31 = 10.18$

This gives $f_{\text{SLED}} \sim 10^{10}$, but the thermal limit at $J = 2$ should be $J^2 = 4$.

The coefficients listed in Extended Data Table 1 appear to differ significantly from those in the original SLED parameterization of Narayanan & Krumholz (2014), from which Equation 6 is derived. This raises the possibility of either an unintended transcription/normalization change, or a different definition of fSLED than the one plotted in Fig. 3. Clarifying how these coefficients are related to the physical CO line ratios (and whether a unit conversion or rescaling is applied) would be needed for clarity.

2) The adoption of a fixed $\alpha_{\text{CO}} = 3.6 \text{ Msun}/(\text{K km/s pc}^2)$ is a critical assumption underlying all Ω_{H_2} inferences. This corresponds to a MW-like conversion factor, but α_{CO} varies with ISM conditions. Could evolution in α_{CO} (e.g., lower values in compact high-SFR systems, or higher values in lower-metallicity galaxies) help explain part of the 2x discrepancy relative to galaxy surveys? Because this assumption sets the absolute normalization of Ω_{H_2} , clarifying its impact is essential to assess how robust the claimed factor-of-two is.

3) The reported 6.6sigma CO detection appears to be a parameter-level significance (i.e., the posterior for a CO amplitude parameter excludes zero within the full 21-parameter model). If so, it would strengthen the claim to also demonstrate a model-level preference by showing that a continuum-only model is statistically disfavored. A Bayes factor or AIC/BIC comparison between continuum-only and continuum+CO(+[CII]) frameworks would clarify that the data actively require a line component rather than merely accommodate one under the assumed parameterization.

Originality and significance:

The work is technically and conceptually original. It extends intensity mapping into a regime where only continuum backgrounds had previously been recovered and demonstrates that legacy broadband maps retain sufficient line information to reconstruct redshift-resolved backgrounds once cross-correlated with tracers. The results are of high interest to multiple communities: galaxy evolution, cosmology, LIM experimental design, and CMB-foreground mitigation.

Data & methodology:

The tomographic cross-correlation method looks robust, the data sets are of high quality, and the separation of continuum and lines via a structured forward model seems appropriate. The author made the data vector and covariance matrix publicly available, and the code is available upon request.

Appropriate use of statistics and treatment of uncertainties:

MCMC posteriors appear to be used correctly, and error bars correspond to 68% credible intervals. The null test (randomizing the line–redshift mapping) is a meaningful validation step, but it tests only the geometric correspondence, not the preference for a line model over a continuum-only model. A Bayes factor or AIC/BIC comparison would address this. In addition, because the depletion time is a ratio of two densities inferred within the same tomographic framework and therefore not statistically independent, it would be helpful to clarify whether their covariance is propagated when computing uncertainty bands on the depletion time, or to note this limitation explicitly.

Conclusions:

The conclusions are broadly supported by the analysis, provided the methodological clarifications and statistical expansions above are addressed. Some interpretative claims should be softened to reflect their dependence on model assumptions, particularly the choice of α_{CO} . The detection of CO and [CII] backgrounds is convincing, but the downstream astrophysical interpretation becomes assumption-dependent and would benefit from clearer qualification.

Suggested improvements:

- Make sure Equation 6, Extended Data Table 1, and Figure 3a are consistent.

- Consider adding a paragraph in Methods discussing α_{CO} uncertainties. If possible, also show how Ω_{H_2} results would change for reasonable variations (e.g., factor of 2).

- If feasible, provide a comparison (Bayes factor/AIC/BIC) between a continuum-only model and the continuum+lines model, to demonstrate that the inclusion of CO is statistically required by the data rather than simply accommodated within the chosen parameterization.

- Clarify whether the covariance between ρ_{H2} and ρ_{SFR} is propagated when computing uncertainty bands on the depletion time, or note this limitation explicitly.

Minor textual suggestions:

Line 28 – “more representative populations that dominate the global budget” is overstated unless quantified.

Line 32 – The companion paper is mentioned but not summarized; add one short descriptive clause so the reader understands what was accomplished there.

Line 37 – “large” is unnecessary and can be removed without loss of meaning.

Line 40 – The explanation “rather than their own emission or shot noise” is too compressed; consider briefly clarifying.

Line 44 – “directly” is misleading, since emissivity remains model-dependent; suggestion: “empirically”.

Line 45 – Add a short definition of emissivity at its first appearance for accessibility.

Line 62 – “unambiguous” reads stronger than the current evidence presentation; consider “high-significance” detection instead.

Line 92 – The phrase “direct Ω_{H2} measurement” overstates the level of inference given fixed α_{CO} .

Line 109 – A bit more detail on the simulations used for benchmarking and how their assumptions map onto your recovered emissivities would strengthen the comparison.

Line 157 – For [CII], given the large uncertainty range, “consistent with” is more accurate than “agrees with.”

Line 170 – This paragraph reads more like part of Results than Conclusion; consider relocating or adjusting tone.

Line 193 – “rather than uncertain forecasts” is too vague; all analyses contain uncertainties. Clarify what specific improvement this work offers (empirical constraint vs. model-dependent inference).

Line 195 – The phrasing “to realize its long-term cosmological promise” is vague. Consider replacing with a more concrete description of the role of these measurements (e.g., “to support forthcoming LIM-based cosmological inference”).

Line 219 – Confirm consistency of redshift coverage (text switches between $z < 4$ and $z < 4.2$).

Line 232 – The equation here deserves to be presented as a stand-alone displayed equation for readability.

Line 242 – Please provide a citation for the expectation referenced here.

Extended Data Fig. 3 – The tick labels are very hard to read at the current size. If space prevents enlarging the full panel, please consider increasing font size or adjusting layout to improve legibility.

Recommendation:

Publish after moderate revisions. The manuscript is conceptually strong and impactful. The requested clarifications and statistical/systematic documentation will strengthen its robustness and accessibility.

References:

Citations appear appropriate and comprehensive.

Clarity and context:

The manuscript is clearly written, the figures are well-presented, and the overall presentation is appropriate for Nature Astronomy’s broad readership.

Inflammatory / DEI concerns:

None.

Expertise boundaries:

I reviewed the cosmological/astrophysical interpretation, cross-correlation methodology, and statistical framework. My assessment of reproducibility is based on the methodology as described, not on re-execution of the pipeline.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #1

(Comments for the Author)

I want to thank the author for responding to my questions and comments. I am satisfied with the changes made to the manuscript. I want to congratulate the author once again for this nice work.

(Remarks on code availability)

I have not reviewed the code. My assessment is based on the methodology presented in the paper.

Reviewer #2

(Comments for the Author)

The authors have been highly responsive in addressing the concerns from my first review. As noted previously, this work makes an important contribution to line-intensity mapping and has significant implications for galaxy evolution, ISM physics, and cosmology. I now recommend the manuscript for publication.

Major Revisions:

CO SLED equation: corrected a transcription error in Equation (6) regarding the inconsistency with Extended Data Table 1 and Figure 3a

alpha_CO uncertainty: a dedicated section now discusses alpha_CO caveats comprehensively

Statistical methodology: the expanded "model-level preference" section provides physical justification for not using Bayes factors. The arguments presented are compelling and appropriate

Systematic uncertainties: added discussion of specific systematics (tSZ, radio background)

Minor points: all satisfactorily addressed

Recommendation: accept for publication

(Remarks on code availability)

made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Referees

We sincerely thank both referees for their careful reading and constructive comments, which have substantially improved the clarity and robustness of the manuscript. We have taken all comments into account and revised both the main text and Methods accordingly. The paper title has been adjusted to follow the Nature Astronomy guideline of no active verb. The revised manuscript incorporates the requested clarifications on the astrophysical interpretation, model assumptions, and statistical framing.

Summary of updates in the revised manuscript:

- added a Methods section to discuss α_{CO} caveats and impacts
- clarified that factor-of-two higher Ω_{H2} compared to galaxy surveys is a relative statement, varying α_{CO} changes the absolute normalizations but not the ratio
- checked and made consistent the interpretation wording throughout, used “direct” or “empirical” for ρ_{line} and “infer” for Ω_{H2}
- added a Methods section on line equivalent width, with a new supporting figure (Extended Data Fig. 6).
- added an intensity version of the T_{b} plot (Extended Data Fig. 5)
- minor readjustment of figure and text order, with small wording revisions throughout the manuscript to improve the overall presentation
- updated the Zenodo release with luminosity density & intensity monopole posteriors.

No changes were made to the data, tomographic measurements, model, or results. Item-by-item responses are provided below, with the original referee comments shown in blue. Figure labels cited below follow the updated ordering in the revised manuscript.

Reviewer #1 (Comments for the Author):

I enjoyed reading the paper. I want to congratulate the authors on this work. Following is my assessment with some questions and comments:

1. Summary:

The paper makes the claims of reporting the first detections of the cosmic CO background aggregate intensity (monopole) at 6.6σ and the cosmic [CII] background at 3.0σ , spanning $0 < z < 4$. The analysis is based on an application of tomographic clustering of diffuse intensity maps (Planck, Herschel, IRAS) cross-correlated with spectroscopic reference galaxies and quasars from the SDSS survey.

2. Originality and Significance:

The work represents an important advance by providing the first empirical constraint on the mean intensity (ρ_{CO}) of the CO background. This approach bypasses the complex, model-dependent integration of the CO luminosity function needed by previous power spectrum measurements (e.g., Keating et al. 2020), offering a crucial normalization factor for the cosmic molecular gas reservoir. The finding that the derived Ω_{H_2} is about twice that of galaxy surveys will directly influence models of galaxy evolution. If the robustness of these results is established, it provides an anchor point for the molecular gas content of the universe, offering a benchmark against which all future theoretical models (e.g., IllustrisTNG, EAGLE) and LIM experiments can be validated. The resulting empirical constraints on the Kennicutt-Schmidt law across cosmic time are important too.

3. Critical Assessment and Comments:

I think the work is largely convincing but is dependent on a chain of non-trivial model assumptions that obscure the physical interpretation of the final astrophysical quantities (Ω_{H_2} , Σ_{SFR}). To fully strengthen the conclusions and make the results more robust, further clarifications are required in the following areas:

MAJOR COMMENTS:

i) Model Dependence and α_{CO} (claims of a "Direct Measurement"):

(Lines 85, 92): The manuscript must clearly distinguish between a 'measurement' and an 'inference' based on model assumptions. The final determination of the molecular gas density Ω_{H_2} is an inferred value, not a "direct measurement". It relies on applying a fixed, non-universal α_{CO} conversion factor. This places the inherent process of this paper in the same spirit as other LIM results (e.g., Breyse et al. 2021) which estimate ρ_{H_2} based on CO intensity, modulated by conversion assumptions. Perhaps the text must be rephrased to emphasize that this is the first unbiased constraint on ρ_{CO} monopole, and the resulting Ω_{H_2} is an inference using the assumed α_{CO} .

The wording has been revised as suggested.

(Lines 95-97): The claim that Ω_{H_2} is twice that of galaxy surveys significantly challenges the models of galaxy evolution. It relies on this same assumption about α_{CO} . α_{CO} is known to vary significantly with

metallicity and gas dynamics (e.g. Bolatto et al. 2013), and the aggregate cosmic population averages over these evolving properties. As a result, this comparison is potentially biased if the galaxy surveys targeted high-metallicity galaxies (low α_{CO}) while the cosmic average requires a different, possibly higher, α_{CO} due to the inclusion of metal-poor populations (Tacconi et al. 2020). It would be good to see an explicit discussion on how the factor-of-two discrepancy would change if a realistic, redshift-dependent $\alpha_{\text{CO}}(z)$ (e.g., scaling with average cosmic metallicity) were adopted. Optionally, the authors can quantify the systematic uncertainty on Ω_{H_2} if α_{CO} were allowed to vary within a conservative, physically plausible range (e.g., 0.8 to 5.0) across $0 < z < 4$. The text must be rephrased to state that the value is inferred via an assumed α_{CO} value.

To clarify our intent, the factor-of-two statement is not presented as a tension with the broader picture of galaxy evolution. Instead, it reflects incompleteness in current high-redshift CO galaxy surveys (Riechers+19; Decarli+19,+20; Lenkic+20; Boogaard+23), measured directly in terms of p_{CO} and conventionally expressed as Ω_{H_2} in those works for physical interpretation. These studies adopt the same Milky Way α_{CO} for cross-comparison, and we follow this convention for consistency rather than to assert a specific value. Because current CO surveys do not yet constrain the faint end of the luminosity function (LF) at high redshift, and in several cases do not sample the knee, they explicitly acknowledge incompleteness (see the main LF plots in aforementioned references). Our higher p_{CO} is therefore best viewed as completing that budget rather than contradicting existing measurements, highlighting LIM as a complementary approach that is complete to lowest surface brightness fluctuations and does not require LF extrapolation.

The main text has been updated accordingly and now states the factor-of-two comparison directly in terms of p_{CO} (1–0). We also appreciate the referee’s emphasis on the uncertainty in α_{CO} . A dedicated section in Methods now summarizes known caveats. As noted there, varying α_{CO} changes the absolute normalization but does not affect the factor-of-two comparison as the same assumption is applied in both the galaxy-surveys and LIM frameworks.

ii) Low- z CII-Line Degeneracy:

(Equation 10-12): The [CII] is unconstrained at $z < 0.4$ due to a gap in spectral coverage. This means the [CII] amplitude at $z=0$ ($R_{\text{[CII]},0}^{100}$) is poorly constrained by the data. The fit from authors uses an evolution parameter,

$C_{\text{[CII]}}^{100}$, which dictates how the [CII] line-to-continuum ratio changes with z (Equation 10-12). This parameter is weakly constrained by the data, reflected by its large posterior error (Extended Data Table 2). The model determines the [CII] cosmic history by simultaneously solving for this $z=0$ anchor ($R_{\text{[CII],0}}^{100}$) and its evolution ($C_{\text{[CII]}}^{100}$) over the range where it is observed ($z > 0.4$). Since the data cannot pin down the $z=0$ anchor point, the fit must extrapolate the CIB continuum (which defines the [CII] baseline) and the evolution of the line down to $z=0$. This is difficult because the weakly constrained ($C_{\text{[CII]}}^{100}$) parameter directly governs the slope of the [CII] history (Fig. 1b). If the model inaccurately estimates the [CII] normalization at $z=0$ (due to weak constraints), this error is carried forward and propagates across all redshifts through the evolution parameter, $C_{\text{[CII]}}^{100}$. A small error in the assumed baseline or slope can significantly change the inferred ρ_{CII} at the peak star formation epoch ($z \sim 2$). The sensitivity is further complicated because at low frequencies/low z , CO lines are most prominent (Fig. 4). Although the CO and [CII] are distinct lines, the spectral fitting process is a complex multi-parameter process; if one part of the low- z spectrum is uncertain due to [CII] extrapolation, it could subtly influence the broader continuum or other co-existing line components.

Have any robustness tests been done to check how sensitive are the high-redshift ρ_{CII} results are to the extrapolation of the CIB continuum model at low frequencies? Have the authors performed a fit using some alternative, motivated prior on the [CII] evolution, such as a fixed (non-evolving) ratio (e.g. $C_{\text{[CII]}}^{100} = 0$, enforcing a flat relationship with the continuum), or a redshift-anchored power law form for $R_{\text{[CII]}}(z)$. If the peak $\rho_{\text{[CII]}}$ result at $z \sim 2$ (where coverage is complete) is truly driven by the data, it should remain stable regardless of the form used for the low- z extrapolation. Forcing $C_{\text{[CII]}}^{100} = 0$ can test whether the weak prior $N(0, 2)$ can artificially steer the high- z results. These tests are just suggestions, but I think the authors should comment on the robustness of this method since it is an important part of claiming the detection and they would add significant credibility to the detection.

We understand the referee's concern, and think the key is to clarify a fundamental behavior of MCMC. In a Bayesian model-fitting framework, the data constrain, in our case, $p_{\text{line}}(z)$ given the allowed family of shapes in redshift evolution described by Equation 13. Since this equation is a linear function in $\log(1+z)$, changing the reference redshift from $z = 0$ to any other value, even $z > 4.2$, via a simple linear transformation permits the same family of $p_{\text{line}}(z)$ curves and thus yields identical posteriors. One direct example is our implementation for $\log \Sigma_{\text{SFR}}$, where the anchor is set to $z = 2$ via a simple linear parameter

transformation. Strictly speaking, this is the same functional form; the $z = 2$ anchor is introduced only to test Gaussian priors directly at that redshift. For all other parameters without informative priors (μ , β , R_{CII}_{100} , R_{CO}), the choice of reference redshift is mathematically arbitrary. Put differently, if a quantity such as $p_{\text{line}}(z)$ is composite (anchor + slope), the MCMC fit constrains the composite rather than individual parameters. The R_{CII} behavior supports this point. The lack of data at $z < 0.4$ naturally broadens the posterior there because, under the same family of possible $p_{\text{line}}(z)$ curves, the high-redshift data can only indirectly constrain $z < 0.4$. Parameterizing the anchor at $z = 0$ does not make the $z = 0$ posterior tighter, consistent with our expectation for MCMC.

On the question of possible CO-[CII] entanglement, we note that the parameterization itself would not induce coupling (only covariance already in the data vector could). This expectation follows from the fact that the two line families are well separated in frequency, and [CII] is modeled as a delta function, which is a maximally “local” spectral feature. The posterior corner plot in Extended Data Fig. 3 confirms this: the four CO parameters and two [CII] parameters show no significant covariance between the two groups, with correlations appearing only within each group (as expected).

iii) CO SLED Expectation:

(Line 242): The authors claim that the SLED should be “representative of the latter [main-sequence galaxies]”. This is an assumption that risks biasing the inferred Σ_{SFR} . The aggregate signal includes all types of galaxies including the luminous starbursts which have highly excited SLEDs. This could be an important caveat because the CO SLED model is calibrated based on a small sample of CO-detected star-forming galaxies at $z=1-3$. This risks introducing a bias based on the properties of those typically bright, resolved sources. Is this a worry? The authors should justify this statement with some prior reference or through some expectations for contribution of main-sequence vs. starburst galaxies to the overall ρ_{TIR} and ρ_{CO} . Will the results on Σ_{SFR} and the steepness of the Kennicutt-Schmidt (K-S) law ($N \approx 1.5$) change significantly based on the scatter in the observed galaxy SLEDs and/or specific choice of the CO SLED function? I don’t think a detailed discussion is required here, just some small justification should suffice.

We clarify that this sentence was intended as an expectation rather than a claim. We cite the population model of Béthermin+22, but the underlying input or target

in such models remains empirical: the bright SMG number counts are low, so their abundance is too small to dominate the global budget. In addition, because CO emission does not necessarily require dust, UV-bright star-forming galaxies with faint IR also contribute, further diluting the relative importance of extreme populations such as SMGs (whose contribution to CO is smaller than to the CIB).

We view our SLED treatment as a strength of this work. Previous LIM studies (e.g., Keating+16; Keating+20; Breysse+22) had to fix the full SLED, whereas our 11-band redshift tomography allows us to constrain the excitation level and its evolution more empirically from the data. Some restriction is still needed given the coarse broadband spectral resolution, so we adopt a family of physically motivated SLED shapes. The recalibration with BzK and CO-selected star-forming galaxies addresses the issue that high-J lines fall too steeply in the original theoretical SLEDs. The resulting SLED family, although calibrated at moderate excitation, can still reproduce a broad range of scenarios from Milky Way-like low excitation to SMG-like high excitation. With this flexibility, the data drive the cosmic SLED posterior toward lower excitation than SMG-like even at the highest redshifts, supporting the expectation above. The best-fit cosmic Σ_{SFR} lies slightly below BzK despite using a weak prior at $z = 2$.

The text describing the SLED framework has been substantially revised.

Because SMG-like excitation is not supported by the data, we did not assess its impact on the inferred K-S slope.

MINOR COMMENTS:

iv) K-S Law Surface vs. Volume Densities:

(Lines 131-140): The analysis linking ρ_{H_2} and ρ_{SFR} (volume) to derive t_{dep} , and then using t_{dep} to calculate Σ_{H_2} and Σ_{SFR} (surface), relies on the assumption that the cosmic-mean depletion time is equal to the average galactic depletion time. This leap from a volume average to a surface average is a central, and slightly non-trivial assumption. It would great if the authors could state this core assumption in the text.

A sentence describing this assumption has been added.

Along the same lines, I have two more curiosities. These following two are optional and a discussion is not needed for the paper:

a) results show that Σ_{SFR} rises with redshift (Fig. 3b). Given that t_{dep} decreases with redshift (Fig. 3c), what is the physical driver here? Does the average star-forming galaxy at high z simply have a smaller star-forming disk size for a fixed stellar mass, or is this primarily a consequence of higher gas fractions/turbulent pressures?

These are open questions for the community, and we have tried to be careful not to over-interpret our results. For $N \approx 1.5$, our interpretation that dense ISM is disproportionately more star-forming is a direct, face-value statement. The mention of enhanced gravitational cloud collapse reflects the classic picture: under an assumption of constant disk height, one could predict an even steeper slope $N \approx 2$, although that assumption may or may not be realistic. This steep scaling expectation motivated subsequent work showing that feedback and turbulence together may moderate the slope to ≈ 1.5 . We cite Krumholz & McKee (2005) in this context, but note that this remains an active area of research aimed at understanding the physical origin of the K-S law (e.g., work by P. Hennebelle and collaborators).

b) implications for Semi-Empirical Models: Fig. 2 shows the Ω_{H_2} results are in better agreement with the EAGLE and semi-empirical model predictions than with IllustrisTNG. What specific physical or empirical parameter in the EAGLE or semi-empirical models (e.g., the halo-mass-to-CO luminosity relation, or the CO-to-H₂ conversion) leads to their higher Ω_{H_2} that aligns better with this novel LIM measurement?

Among the models themselves, the predictions are for H₂ rather than CO, so the conversion factor does not enter this comparison. One relevant difference is that IllustrisTNG appears to underproduce cosmic star formation at $z < 3$ (see Fig. 4 in Pillepich+17). There are likely additional contributing factors, several of which are discussed in detail in Popping+19, but they are highly model-specific and therefore beyond the scope of our text.

v) Thermal Sunyaev-Zel'dovich effect (tSZ) and radio contamination:

The Planck maps contain tSZ (100, 143, and 353 GHz; although smaller than CIB at 353 GHz) and radio emission from extragalactic galaxies (mostly at 100 and 143 GHz). Since these emissions also trace the large-scale structure, they should be correlated with the galaxies and quasars from the SDSS galaxy surveys used here. While the companion paper mentions fitting for a separate tSZ term, no such thing has been mentioned here. The authors should clarify

how this has been dealt with. Also, the authors should specify if they expect the radio contamination to be negligible if it has been neglected.

This information has been added in the first section of Methods: “As described in the companion work, a small thermal Sunyaev-Zeldovich distortion from hot gas in the cosmic web, detected jointly at low frequencies, has already been removed from the CIB (dI/dz) data vector. Radio free-free and synchrotron emission are undetected even in our lowest-frequency band (100 GHz), with their inclusion yielding only upper limits. We therefore do not include them in this work.”

In the Monopole section of Methods, we additionally show that the SIDES model predicts a consistent behavior in which radio backgrounds are negligible at 100 GHz, with contributions beginning to rise only below that frequency.

vi) Neglecting Higher-J CO lines and Interlopers:

(Lines 55-56): The model explicitly stops fitting the CO ladder at $J=9-8$. High-J transitions ($J \geq 10$) are powerful tracers of the dense, highly excited gas that characterizes starbursts (e.g. Greve et al. 2014) and should contribute to the observed bands used here. Ignoring these ignores a fraction of the total luminosity and potentially changes the inferred Σ_{SFR} based on SLED modeling.

At the same time, the paper neglects other bright atomic lines ([OIII], [CII]) and dense gas tracers (HCN, HCO+) that correlate with star formation and fall within the observational windows. If these are expected to be negligible, it would make the paper more thorough to mention this justification for the exclusion of $J \geq 10$ lines as well as other interloper lines.

These lines are indeed too faint to be detectable given the precision of our tomographic data vector, and we have added a note in the text. At high frequency we reach only $\approx 3\sigma$ even for [CII], which is the brightest line, so fainter lines would be harder to detect. For CO, the high-J contribution in our result is already negligible in luminosity density, and even more so in equivalent width as the CIB continuum rises toward higher frequencies. For readers interested in forecasting $J \geq 10$, our constraints can be extrapolated, but the SLED calibration (Extended Data Table 1) is not defined beyond $J = 9-8$, so any such extension would require adopting an external assumption or extrapolating from Fig. 4a.

vii) SLED Ceiling:

(Line 251): Provide a literature reference or argument for the fully thermalized ceiling assumption ($f_{\text{SLED}} = J^2$ at $\log \Sigma_{\text{SFR}} = 4$).

We have revised the text to clarify that the J^2 scaling is not introduced ad hoc but reflects the standard thermalized limit of the CO ladder, consistent with the high-excitation behavior in Narayanan & Krumholz (2014) and with level-population theories in ISM textbooks. The ceiling at $\log \Sigma_{\text{SFR}} = 4$ is set semi-arbitrarily to lie well above any realistic starburst ISM and beyond the regime where the SLED is already saturating or nearly thermalized, so the exact choice of the anchor ($\log \Sigma_{\text{SFR}} = 4$) does not impact our cosmic CO results. Nonetheless, having such a ceiling ensures numerical stability when fitting the coefficients in Extended Data Table 1.

viii) CONCERTO Reference:

(Line 365): CONCERTO is not a model. Correct the phrasing to confirm that Ref. 64 (B  thermin et al. 2022) is cited as a simulation/model (a forecast for line emissions) that produces results consistent with the detection, rather than referring to the observational experiment.

Changed CONCERTO to SIDES as suggested.

ix) CO First Detection:

In the abstract, the authors mention that this is the first detection of the cosmic CO background. While the distinction between measuring the monopole (mean intensity) and the power spectrum (shot noise) is key, the term first detection can be ambiguous given previous LIM work like COPSS and mmIME. It would be good to clearly state that this first detection refers specifically to the mean intensity (ρ_{CO}), which is the integral (first moment) of the luminosity function, as opposed to the power spectrum (second moment) detected previously.

As suggested, we have added “mean intensity” in the abstract. Nonetheless, to most readers outside the LIM community, the term “cosmic background,” if not further qualified, naturally refers to the mean intensity or total light (not shot power as detected in previous work).

x) ρ_{SFR} Consistency and Possible Circular Logic:

(Lines 121-123): The dependence of the depletion time (t_{dep}) on $\rho_{\text{SFR}}(z)$ (Line 122), which references the CIB-only companion study, could be problematic because that original model neglected line contributions. While the present joint fit seems to correct the biased CIB parameters (σ_{β} changes from 0.71 to 0.24), the language suggests using the old result from the companion paper. Can the authors please explicitly confirm in the Methods section that the $\rho_{\text{SFR}}(z)$ used for t_{dep} is derived from the corrected CIB continuum parameters obtained from the joint CIB+lines fit performed in this current paper or the old paper?

We have clarified and expanded the justification in the text that the $\rho_{\text{SFR}}(z)$ used for t_{dep} comes from the companion CIB-only analysis, specifically from the smooth form in their Equation 42, with a 10% uncertainty adopted. We retain this choice for reproducibility, as the companion paper provides the full numerical presentation of the SFR history. Equation 42 also includes the modest UV correction, which is not rederived in this line-focused paper.

For t_{dep} , uncertainties remain at the factor-of-two level across redshift, so small line corrections are not significant. For completeness (beyond what is appropriate for the main text), we note that the line contribution to the total CIB energy budget is ≈ 0.003 from [CII] and ≈ 0.0003 from CO, and the σ_{β} shift from ≈ 0.71 to ≈ 0.24 alters the total infrared luminosity density at a similarly small level. These two effects largely offset, making the net change in the total IR even smaller. One may view this cancellation as a consequence of the total IR being an observable rather than a derived physical quantity, and already well sampled by the tomographic data vector. Thus, interpreting it via continuum-only or continuum+lines has limited practical impact, as both must fit the same underlying measurements.

Another reason that we choose not to be overly precise is that the IR-to-SFR conversion factor from the original Kennicutt calibration, or more recent variants, carries a much larger systematic uncertainty than any line contribution, and CO is not always removed from the dust continuum when those calibrations were derived.

Congratulations on the work once again,

Reviewer #1 (Remarks on code availability):

NA

Reviewer #2 (Comments for the Author):

Key Results:

This manuscript presents the first tomographic detection of the cosmic CO and [CII] backgrounds over $0 < z < 4$ using cross-correlations between broadband far-infrared/millimeter intensity maps (Planck, Herschel, IRAS) and spectroscopic tracers from SDSS. The authors recover redshift-resolved, bias-weighted emissivities and forward-model the continuum and line components using a 21-parameter framework combining a graybody CIB model with a CO SLED and a [CII] component. They infer a molecular gas density approximately two times higher than galaxy survey estimates at $1 < z < 2.5$, a depletion time of 1 Gyr, and a persistent large-scale Kennicutt–Schmidt relation. These results provide an important empirical baseline for future LIM experiments.

Validity:

I do not find problems that would preclude publication, but I do identify important points that require clarification, correction or expansion before the results can be considered robust:

1) There appears to be a significant inconsistency between Equation (6), Extended Data Table 1, and Figure 3a. This may be a definitional or scaling issue, but it requires clarification or correction.

For example:

For $J=2-1$, at $\log \Sigma_{\text{SFR}} = 0$ (i.e., 1 Msun/yr/kpc^2):

$\log f_{\text{SLED}} = 0.3 \times (0 - (-2))^{1.95} + 0.31 = 1.47$

So: $I(2-1)/I(1-0) = 10^{1.47} \approx 30$

But Fig 3a shows at $J_{\text{up}}=2$ the cosmic ratios are $\sim 0.5\text{--}3$.

The text also states f_{SLED} should reach J^2 at $\log \Sigma(\text{SFR}) = 4$.

For $J=2-1$ at this $\Sigma(\text{SFR})$ value:

$\log f_{\text{SLED}} = 0.3 \times (4 - (-2))^{1.95} + 0.31 = 10.18$

This gives $f_{\text{SLED}} \approx 10^{10}$, but the thermal limit at $J = 2$ should be $J^2 = 4$.

The coefficients listed in Extended Data Table 1 appear to differ significantly from those in the original SLED parameterization of Narayanan & Krumholz (2014), from which Equation 6 is derived. This raises the possibility of either an unintended transcription/normalization change, or a different definition of f_{SLED} than the one plotted in Fig. 3. Clarifying how these coefficients are related to the physical CO line ratios (and whether a unit conversion or rescaling is applied) would be needed for clarity.

We thank the referee for noting this inconsistency, which revealed an unintended mistake that the written form of the equation did not match the numerically implemented version. We have corrected the printed equation accordingly. All analyses, results, figures (including Fig. 4a), and the coefficients in Extended Data Table 1 were already based on the correct functional form, and our MCMC code submitted to the Editor previously remains unchanged.

As now stated in the revised manuscript, the form actually used is

$$\log f_{\text{SLED}} = A (x^B) / (1 + x^B) + C, \text{ with } x = \log \Sigma_{\text{SFR}} - (-2),$$

whereas the original Narayanan & Krumholz (2014; NK14) form was

$$\log f_{\text{SLED}} = A (x^B) + C.$$

Using this stabilized form, we then recalibrated the coefficients with multi-J observations of typical star-forming galaxies (Daddi+15; Boogaard+20). The refinement matters primarily at moderate Σ_{SFR} , where there was a lack of empirical comparison in NK14 due to limited observational data available at the time. Two further notes:

- without recalibration, our modified function can still reproduce the original NK14 SLEDs (with different A, B, C), meaning the two parameterizations could define the same SLED family, given the flexibility allowed by free coefficients for each J;
- with recalibration, new A, B, C coefficient are determined; the high-J decline becomes less steep, bringing the modeled SLEDs into closer agreement with the newer measurements in this Σ_{SFR} regime.

As shown in Fig. 4a, CO(3–2) galaxies from Boogaard+20 at $\Sigma_{\text{SFR}} \approx 2 \text{ M}_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ are matched across $J = 1\text{--}8$ by our recalibrated family, whereas no single Σ_{SFR} could reproduce the full ladder using the original NK14 parameterization (their Fig. 7). Conversely, NK14 already performs well for high- Σ_{SFR} starbursts (their Fig. 8), where both formulations approach the thermalized J^2 limit. The difference thus reflects new leverage at lower Σ_{SFR} ,

which was not tested against observations when NK14 was developed.

The CO excitation section in Methods has been substantially expanded to clarify the motivation for the modified form, the recalibration procedure, and its impact on the inferred CO excitation in this work.

2) The adoption of a fixed $\alpha_{\text{CO}} = 3.6 \text{ Msun}/(\text{K km/s pc}^2)$ is a critical assumption underlying all Ω_{H2} inferences. This corresponds to a MW-like conversion factor, but α_{CO} varies with ISM conditions. Could evolution in α_{CO} (e.g., lower values in compact high-SFR systems, or higher values in lower-metallicity galaxies) help explain part of the 2x discrepancy relative to galaxy surveys? Because this assumption sets the absolute normalization of Ω_{H2} , clarifying its impact is essential to assess how robust the claimed factor-of-two is.

A similar point was also raised by Reviewer #1, and part of our response is provided above. As clarified in the revised text, we now state explicitly that the factor-of-two excess relative to galaxy-survey estimates refers to $\rho_{\text{CO}(1-0)}$, and the interpretation is survey incompleteness rather than a physical tension. This incompleteness was already acknowledged in the cited works but could not be quantitatively evaluated without a LIM-based total budget prior to our result here.

The uncertainty in α_{CO} affects only the absolute normalization in Fig. 3: most observational ρ_{CO} values would shift together under a different α_{CO} assumption, while keeping the ratios unchanged (with two noted exceptions in the text, which adopt different α_{CO} assumptions).

We have added a dedicated section in Methods to discuss the role and caveats of α_{CO} .

3) The reported 6.6sigma CO detection appears to be a parameter-level significance (i.e., the posterior for a CO amplitude parameter excludes zero within the full 21-parameter model). If so, it would strengthen the claim to also demonstrate a model-level preference by showing that a continuum-only model is statistically disfavored. A Bayes factor or AIC/BIC comparison between continuum-only and continuum+CO(+[CII]) frameworks would clarify that the data actively require a line component rather than merely accommodate one under the assumed parameterization.

Yes, the quoted 6.6σ reflects the posterior exclusion of zero for ρ_{CO} , and ρ_{CO} is a derived quantity from the full 21-parameter fit.

Regarding model-level preference, we have expanded the discussion in the revised section now titled “Model-level preference.” We think AIC/BIC might not be directly informative in this case because the lines are faint relative to the CIB continuum ($\approx 0.3\%$ for [CII] and $\approx 0.03\%$ for CO), so the log-likelihood in AIC/BIC would be largely driven by the continuum.

Our main arguments for model selection are therefore more physical ones. Without lines, the continuum-only fit requires an unphysically large $\sigma_\beta \approx 0.71$ to absorb the CO ladder excess, whereas including CO+[CII] (particularly CO) yields a smaller σ_β that is more physically plausible.

Moreover, CO and [CII] emission at these rest-frame frequencies is fixed by molecular transition physics, so including them defines the appropriate model space. In this setting, the usual “look-elsewhere effect” does not apply, as we are not searching blindly over unknown frequencies; the line frequencies are predetermined, and within this physical model space the data favor positive amplitudes. These clarifications are now more explicitly stated in the revised text.

Originality and significance:

The work is technically and conceptually original. It extends intensity mapping into a regime where only continuum backgrounds had previously been recovered and demonstrates that legacy broadband maps retain sufficient line information to reconstruct redshift-resolved backgrounds once cross-correlated with tracers. The results are of high interest to multiple communities: galaxy evolution, cosmology, LIM experimental design, and CMB-foreground mitigation.

Data & methodology:

The tomographic cross-correlation method looks robust, the data sets are of high quality, and the separation of continuum and lines via a structured forward model seems appropriate. The author made the data vector and covariance matrix publicly available, and the code is available upon request.

Appropriate use of statistics and treatment of uncertainties:

MCMC posteriors appear to be used correctly, and error bars correspond to 68% credible intervals. The null test (randomizing the line–redshift mapping) is a meaningful validation step, but it tests only the geometric correspondence, not the preference for a line model over a continuum-only model. A Bayes factor or AIC/BIC comparison would address this. In addition, because the depletion time is a ratio of two densities inferred within the same tomographic framework and therefore not statistically independent, it would be helpful to clarify whether their covariance is propagated when computing uncertainty bands on the depletion time, or to note this limitation explicitly.

Model-level preference is addressed above and in the expanded section in Methods.

Regarding depletion time, $t_{\text{dep}} = \rho_{\text{H2}} / \rho_{\text{SFR}}$, the statistical uncertainty is dominated by ρ_{H2} . As clarified in response to Reviewer #1, we use the smooth $\rho_{\text{SFR}}(z)$ form from Equation 42 of the companion CIB-only paper with a 10% uncertainty adopted to match the stated 0.04 dex in the abstract of that work. In this setup, covariance is fully propagated up to ρ_{H2} (which depends on the CO amplitude covarying with the SLED, and carries weak but non-zero covariance with the continuum), while ρ_{SFR} is treated as an external input with its stated uncertainty. In practice, $\rho_{\text{H2}}(z)$ is computed for every MCMC realization of the 21-parameter posterior (given an assumed α_{CO}) on dense redshift grid, and the spread in ρ_{H2} posterior dominates the uncertainty in t_{dep} . The revised text now states this more clearly.

The K-S relation in Fig. 4d does propagate the full covariance: each MCMC draw is mapped into $(\Sigma_{\text{SFR}}, \Sigma_{\text{H2}})$, and the resulting 1- σ contour is a tilted ellipse, reflecting the joint posterior rather than independent error bars.

Conclusions:

The conclusions are broadly supported by the analysis, provided the methodological clarifications and statistical expansions above are addressed. Some interpretative claims should be softened to reflect their dependence on model assumptions, particularly the choice of α_{CO} . The detection of CO and [CII] backgrounds is convincing, but the downstream astrophysical interpretation becomes assumption-dependent and would benefit from clearer qualification.

Suggested improvements:

- Make sure Equation 6, Extended Data Table 1, and Figure 3a are consistent.
- Consider adding a paragraph in Methods discussing α_{CO} uncertainties. If possible, also show how Ω_{H2} results would change for reasonable variations (e.g., factor of 2).
- If feasible, provide a comparison (Bayes factor/AIC/BIC) between a continuum-only model and the continuum+lines model, to demonstrate that the inclusion of CO is statistically required by the data rather than simply accommodated within the chosen parameterization.
- Clarify whether the covariance between ρ_{H2} and ρ_{SFR} is propagated when computing uncertainty bands on the depletion time, or note this limitation explicitly.

Minor textual suggestions:

Line 28 – “more representative populations that dominate the global budget” is overstated unless quantified.

We expanded the text to more explicitly describe the known limitations and observational challenges of CO and [CII] galaxy surveys.

Line 32 – The companion paper is mentioned but not summarized; add one short descriptive clause so the reader understands what was accomplished there.

We appreciate the suggestion and have expanded the text to provide a summary of the companion work.

Line 37 – “large” is unnecessary and can be removed without loss of meaning.

Revised.

Line 40 – The explanation “rather than their own emission or shot noise” is too compressed; consider briefly clarifying.

As suggested, we expanded the text and clarified that shot noise refers to fluctuations dominated by discrete bright sources.

Line 44 – “directly” is misleading, since emissivity remains model-dependent; suggestion: “empirically”.

Removed “directly”.

Line 45 – Add a short definition of emissivity at its first appearance for accessibility.

Definition added.

Line 62 – “unambiguous” reads stronger than the current evidence presentation; consider “high-significance” detection instead.

Revised to “We detect a statistically significant 6.6σ cosmic CO signal”

Line 92 – The phrase “direct Ω_{H_2} measurement” overstates the level of inference given fixed α_{CO} .

We revised the sentence to make clear that Ω_{H_2} is derived from ρ_{CO} , modulo α_{CO} .

Line 109 – A bit more detail on the simulations used for benchmarking and how their assumptions map onto your recovered emissivities would strengthen the comparison.

We expanded the description. By explicitly noting that these models predict H_2 directly, without requiring an assumed α_{CO} , the text now makes clear that detailed comparisons between simulations and observations carry an additional systematic uncertainty. This is also why we do not push the simulation–data comparison as far as the observational comparison between LIM and galaxy surveys, where ρ_{CO} is directly measured in both cases and the α_{CO} dependence cancels in the ρ_{H_2} ratios.

Line 157 – For [CII], given the large uncertainty range, “consistent with” is more accurate than “agrees with.”

Revised accordingly.

Line 170 – This paragraph reads more like part of Results than Conclusion; consider relocating or adjusting tone.

Moved this paragraph, together with the associated T_b plot, earlier in the manuscript. We also take this chance to align with Nature Astronomy format guidelines by removing explicit section labels such as “Results” and “Conclusion,” maintaining a continuous narrative throughout.

Line 193 – “rather than uncertain forecasts” is too vague; all analyses contain uncertainties. Clarify what specific improvement this work offers (empirical constraint vs. model-dependent inference).

Revised this sentence to “Future experiments can therefore be designed and optimized using empirical targets rather than a broad span of model predictions.”

Line 195 – The phrasing “to realize its long-term cosmological promise” is vague. Consider replacing with a more concrete description of the role of these measurements (e.g., “to support forthcoming LIM-based cosmological inference”)

We revised the few sentences in this ending part to be more concrete.

Line 219 – Confirm consistency of redshift coverage (text switches between $z < 4$ and $z < 4.2$).

Revised to $z < 4.2$ consistently.

Line 232 – The equation here deserves to be presented as a stand-alone displayed equation for readability.

Revised as suggested.

Line 242 – Please provide a citation for the expectation referenced here.

Reference is added and text expanded.

Extended Data Fig. 3 – The tick labels are very hard to read at the current size. If space prevents enlarging the full panel, please consider increasing font size or adjusting layout to improve legibility.

Tick labels have been enlarged as suggested.

Recommendation:

Publish after moderate revisions. The manuscript is conceptually strong and impactful. The requested clarifications and statistical/systematic documentation will strengthen its robustness and accessibility.

References:

Citations appear appropriate and comprehensive.

Clarity and context:

The manuscript is clearly written, the figures are well-presented, and the overall presentation is appropriate for Nature Astronomy's broad readership.

Inflammatory / DEI concerns:

None.

Expertise boundaries:

I reviewed the cosmological/astrophysical interpretation, cross-correlation methodology, and statistical framework. My assessment of reproducibility is based on the methodology as described, not on re-execution of the pipeline.